Comments on the Second Draft Report: "Assessment of Nutrient Loading and Eutrophication in Barnegat Bay-Little Egg Harbor, NJ in Support of Nutrient Management Planning"

January 2013

EPA Comments

The entire developmental process for each indicator (including detailed basis for the thresholds, and how the index is calculated from the components) must be supplied. While the thresholds for most of the indicators are included (benthic invertebrates is missing), the rationale behind the establishment of these thresholds is still unclear. The threshold values and weightings for the benthic invertebrate indicator must be provided or this indicator removed from the index.

Report what the criteria and results were for the screening and selection of individual indicators. Specifically, what were the criteria for spatial representativeness (data for each segment), temporal representativeness (i.e. seasonal or diurnal data) or data completeness? In addition, some indicators or data appear to have been excluded, for example: water residence time and bay scallops. Summarize and explain why certain indicators or data sets were excluded. This might best be accomplished in tabular format.

The evaluation process for secondary data (NJDEP and any other data sources) still needs to be supplied. This should include a description of how these data were screened for use. In particular, what temporal and spatial coverage criteria were used, how outliers were determined and how they were dealt with, how non-detects and missing values were treated, criteria for ensuring comparable sampling and laboratory methods, etc. Representativeness of the secondary data also should have been evaluated and needs to be stated. This information should be summarized in a table along with source citations for the data.

QA/QC results for the data collected as a result of this project (Rutgers data) must be summarized and reported. For example; results of duplicates and completeness of data should be included (response document stated that this was on pg. 75 but there are no QA results there). The QA/QC results should be given based on the Measurement Quality Objectives (MQOS) stated in the QAPP.

A great deal of information needs to be provided on the use of the Principal Components Analysis (PCA) since it is not a standard approach. The statistical concerns raised in our comments of the first draft of this document, as well as those put forth by NJDEP's statistician dated August 21, 2012 still need to be addressed, including the lack of detail regarding non-detects, zeros, missing data, outliers, weighting, and squaring the eigenvalue. The weaknesses of PCA should be discussed and related to its use in the index. One of the assumptions for PCA is that sets of scores should represent a random sample from the population of interest. Was that the case in this application?

Component 4 of the report only reports the data collected in 2011. This section must also discuss how the validation was done (provide the statistical procedures used) and the results of the validation of the index using the 2011 data.

The final report should include the citations of field and lab methods and QA/QC procedures and any modifications that were made, as this is standard scientific procedure. Report readers should not have to track down additional documents to understand what methods were used.

The response document states that the QAPP is cited in the revised report but we couldn't find it.

Additional data from a separate seagrass project in northern Barnegat Bay were used with this study. Since these were not discussed in the QAPP for the current (NEIWPCC) project, discuss method and data comparability, reference the appropriate QAPP and final report for that study, and describe the process and thresholds for incorporating *Ruppia* into the index.

We have repeatedly asked to see the raw data (SAV and water quality) collected as part of this study and are assured that it will be made available. We would like to see this dataset now rather than at the completion of the project, in order to better understand how it was used. Two weeks prior to the soon-to-be scheduled TAC meeting with the researchers would provide us with enough time for a preliminary review.

Biotic Index Development: The QAPP (pg. 28) stated: "They (biotic data) will be examined and assessed for statistical validity and inclusion in the index development for the 1989 to 2011 period." The actual statistical procedures and outcomes should be included in the report.

Some specific examples of issues which we hope will be clarified once the above "big picture" issues are addressed:

- The benthic invertebrate data from 2001 were not included in Table 3-15. It is unclear whether the benthic invertebrate data were used to calculate the index
- Per Figure 3-2, Loading data are available for the entire time period, yet Table 3-15 does not include this component in the calculation of the overall index. The report must explain why these data were not used.
- Per Figure 3-2 total phosphorus data were available in 1999, yet according to Table 3-15 these data were not used in calculation of the index. The report must explain why these data were not used.
- Per Figure 3-2, there seems to have been data available for Chlorophyll *a*, TSS and Secchi readings in 1997, yet Table 3-15 indicates these data were not used. The report must explain why these data were not used.
- Rutgers was to have collected water quality data during their seagrass sampling efforts. It is unclear whether or not these data were used in calculation of the index. Figures 2-1 through 2-7 only show the NJDEP data.
- Biotic Response, Seagrass: This section stated that seagrass assessments were "adjusted" because of "the uncertainty associated with identifying reference conditions in BB-LEH." Please provide additional detail on how these adjustments were made.
- Bay scallop data: The response to comments state that this data is stored in an archived database that is accessible to the TAC. Please identify the location of this database.
- For Figures 3-14 to 3-16 and 3-23 to 3-24: Please clarify, do these blue lines represent a linear regression line?

NJ DEP Comments

The Department provided extensive comments on the draft report and we believe that many of the response provided were not responsive. Often the response indicated that the comments were beyond the scope of the project or not in the QAPP. At this point, we are limiting our focus

on the most critical portions of the report and that is the selection of thresholds, how the data is summarized to calculate raw scores and how the PCA analysis was used to develop weightings and ultimately the index. We had always hoped to have good spirited discussion with the researchers on their recommended thresholds but that never happened. The spreadsheets developed by EPA for the researchers to complete will go a long way in providing the information we need to evaluate the researchers' recommendation.

The researchers must justify the selected thresholds in Tables 3-3, 3-8, 3-10, 3-11 and 3-12 as there does not appear to be a direct connection to the thresholds identified in the selected references. As an example, the selected thresholds for TN (Table 3-8) is 400ug/l for a score of 13 (poor) while the basis for this threshold is listed at Table 3-6 with then that 0.55mg/l is better than seagrass objectives. It appears that our poor meets Maryland Inland Bays "better than seagrass" objective. The researchers need to document their thresholds and not just say we applied "best profession judgment". There needs to be documentation supporting the decisions.

We also have concerns on how the raw data is summarized and whether the approach used is sound. Appendix 3-4 describes the "lightly summarized" data (mean or median) by year for North, Central and South. We need to understand which stations were associated with which the Northern, Central, and Southern parts of the bay. A map showing the clustering of stations would be helpful or a table indicating the station location and the portion of the bay the stations was associated with. Appendix 3-4 should be modified to include the source for the data, number of stations/samples used to calculate the value a given year and describe the QA review applied to the data before calculating the means/medians. For example, how were "nondetects" handled? How were missing samples handled?

For the water quality indicators, the source is DEP quarterly data. However, the key question going forward is how the data should be summarized as we will have more intensive monitoring. While this may be beyond the scope of this project, it is critical that we understand why annual means/medians were used and whether that's the most appropriate summarization method.

We still have questions concerning the appropriateness of calculating an annual mean/median for parameters that vary seasonally such as DO or temperature. This is not consistent the reference provided at Table 3-5. The DO thresholds identified for Maryland Coastal Bays is based on data from June to August. In addition, the biologically relevant threshold for Maryland Coastal Bays identified greater than 7mg/l as "better than seagrass" objectives. This value is more consistent with 100% saturation in warm, estuarine waters. While the 10mg/l threshold identified in Table 3-8 could only occur in the summer if the waters were superstaturated.

Table 3-12 provides the thresholds for Harmful Algal Blooms. DEP now has equipment to identify and quantify *A. anophagefferns* if the concentrations are expected at levels greater than 10,000 cells. We acknowledge that there may not be data for all years, but we clearly need to understand how this data should be summarized in the future.

In revised report, the conclusion recommends limiting the loading of nitrogen to 1500 kg/km^2/year and phosphorus to 75 kg/km^2/year. This recommendation should be consistent with Table 3-3 and validated against the natural loading expected if the watershed consisted of only forest and wetlands.

At this point in the process, funds may not be available for the researchers to lightly summarize 2011 data and apply the index. However, it is absolutely critical that the researchers document how the data was summarized if we have any intention on applying the index on new data such as 2011/12 or another estuary. Once the summarization process is adequately documented we should be able to duplicate the actual "raw" values for any given year by following their methodology. In addition, we need to understand how to calculate the weighted scores and how this information feeds into the index.

Barnegat Bay Estuary Partnership

As a result of the inconsistent responses to all of the TAC's comments on this report, our evaluation of the revised report remains very mixed. We do note notable improvements to some sections of the report, especially Section 1, where the USGS report has been well-integrated into the report, and Section 5, where some management recommendations have been included. Nonetheless, we have considerable reservations about this report in its present form.

Organization

Though the organization of the report has been improved in places, the methods are inconsistently and, in critical sections, inadequately explained. Our and others' specific comments were inconsistently responded to; some references were not included; these inconsistencies and omissions, in our view, weaken the overall report. Recognition of important data gaps and acknowledgement of other potential explanations of the observations and findings of this study would strengthen the manuscript overall.

Methods

The purpose for including some data and/or methods in the project is not sufficiently clear. The methods remain inconsistently organized among sections and inadequately described within some sections. This is particularly true in the index development (Section 2) and the model validation section (Section 4). The report should more simply and concisely present the methods, including the collection of data, statistical tests used to analyze the data, and identify one or more appropriate citations regarding the use of the chosen methods. Weaknesses in the use of some pressure metrics (total N, which clearly overestimates the available N in estuaries, even lagoonal ones [Seitzinger and Sanders, 1997]) should be acknowledged.

The authors should address the use of many more metrics used for bioindicators in comparison to other models (a point the authors emphasize upfront!). Using some additional metrics has some potential utility, but not all of these bioindicators respond solely to eutrophication. Different indicators may change with other pressures. Other explanations for changes in the indicators are not adequately or consistently addressed. Too many indicators appear forced as eutrophication responses, whereas other causes of changes in indicators (e.g., overfishing on clams) are both reasonable and likely to have occurred in the system over time. For example, the causes of the decline of clams may be different than those factors which keep clam abundance and/or recruitment low. Which indicators appear strongly tied to eutrophication; which appear loosely tied to eutrophication?

In Component 3, the PCA figures have the variables broken by segment; moreover, the indices are reported by segment. However, the table that contains the weightings is for the estuary as a whole. Would you please clarify/justify this approach/inconsistency?

Component 4 does not lay out the appropriate methods or results (the generated indices and the fit of the 2011 data to the eutrophication model). How the model was validated using 2011 data? Component 4 simply results the results of the data collected during 2011.

Results

There are some troubling inconsistencies in this report. Perhaps most alarming, the bay is reported to exhibit "an insidious system-wide decline" yet the study's findings point out that the different segments and components of the system exhibit distinct patterns and trends, some of which may even be improving. This inconsistency may be resolved by more clearly identifying critical data gaps, even those outside of the scope of the QAPP. For example, bacterial production may help explain many observations in the condition of the bay, but microbial production and the ecosystem consequences of changes to the "microbial loop" are not mentioned anywhere in the report.

Contrary to the author's claim, Guo and Psuty 2000 and Guo et al 2004 suggest that the nearshore coastal ocean, which receives considerable secondary sewage effluent discharges, may be serving as a considerable source of nutrients to the system episodically. They show nutrient levels 1-2 orders of magnitude higher periodically; these findings directly contradict statements in the report. The study's conclusions should acknowledge this/other noted observations (e.g., OCNGS generation of chloramines/other N compounds) as data gaps which may explain some of the interesting findings reported in the study.

The authors define eutrophication differently from others. Nixon defines eutrophication as an increase in the delivery of organic matter to an ecosystem, and not as an increase in nutrients and organic matter to an ecosystem. This may seem like semantics, but by this simple change in wording, eutrophication causes eutrophication which may lead to some confusion, especially in readers, in causes and effects of the changes in the bay.

Lastly, the report still includes statements which appear to be opinion because they are not supported adequately by citations or presented evidence.